2. Tournaments

A *tournament* is an oriented graph $T = (V, E)$ such that $(x, x) \notin E$ for all $x \in V$, and for any two vertices $x \neq y$ exactly one of (x, y) and (y, x) belongs to E. That is, each tournament is obtained from a complete graph by orienting its edges. The name tournament is natural, since one can think of the set V as a set of players in which each pair participates in a single match, where $(x, y) \in E$ iff x beats y.

Say that a tournament has the property P_k if for every set of k players there is one who beats them all, i.e., if for any subset $S \subseteq V$ of k players there exists a player $y \notin S$ such that $(y, x) \in E$ for all $x \in S$.

THEOREM 3.1 (Erdős 1963a). *If $n \geq k^2 2^{k+1}$, then there is a tournament of n players that has the property P_k.*

PROOF. Consider a random tournament of n players, i.e., the outcome of every game is determined by the flip of fair coin. For a set S of k players, let A_S be the event that *no $y \notin S$ beats all of S*. Each $y \notin S$ has probability 2^{-k} of beating all of S and there are $n - k$ such possible y, all of whose chances are mutually independent. Hence $\Pr[A_S] = (1 - 2^{-k})^{n-k}$ and

$$\Pr\left[\bigcup A_S\right] \leq \binom{n}{k}(1 - 2^{-k})^{n-k} < \frac{n^k}{k!}\mathrm{e}^{-(n-k)/2^k} \leq n^k \mathrm{e}^{-n/2^k}.$$

If $n \geq k^2 2^{k+1}$, this probability is strictly smaller than 1. Thus, for such an n, with positive probability no event A_S occurs. This means that there is a point in the probability space for which none of the events A_S happens. This point is a tournament T and this tournament has the property P_k. $\square$

3. Universal sets

A set of 0-1 strings of length n is (n, k)-*universal* if, for any subset of k coordinates $S = \{i_1, \ldots, i_k\}$, the projection

$$A\restriction_S := \{(a_{i_1}, \ldots, a_{i_k}) : (a_1, \ldots, a_n) \in A\}$$

of A onto the coordinates in S contains all possible 2^k configurations.

On the other hand, a simple probabilistic argument shows that (n, k)-universal sets of size $k2^k \log_2 n$ exist (note that 2^k is a trivial lower bound).

THEOREM 3.2 (Kleitman–Spencer 1973). *If $\binom{n}{k}2^k(1 - 2^{-k})^r < 1$, then there is an (n, k)-universal set of size r.*

PROOF. Let $\boldsymbol{A}$ be a set of r random 0-1 strings of length n, each entry of which takes values 0 or 1 independently and with equal probability 1/2. For every fixed set S of k coordinates and for every fixed vector $v \in \{0, 1\}^k$,

$$\Pr[v \notin \boldsymbol{A}\restriction_S] = \prod_{a \in \boldsymbol{A}} \Pr[v \neq a\restriction_S] = \prod_{a \in \boldsymbol{A}} \left(1 - 2^{-|S|}\right) = \left(1 - 2^{-k}\right)^r.$$

Since there are only $\binom{n}{k}2^k$ possibilities to choose a pair (S, v), the set $\boldsymbol{A}$ is *not* (n, k)-universal with probability at most $\binom{n}{k}2^k(1 - 2^{-k})^r$, which is strictly smaller than 1. Thus, at least one set A of r vectors must be (n, k)-universal, as claimed. $\square$

4. Covering by bipartite cliques

A *biclique covering* of a graph G is a set $H_1, \dots, H_t$ of its complete bipartite subgraphs such that each edge of G belongs to at least one of these subgraphs. The *weight* of such a covering is the sum $\sum_{i=1}^{t} |V(H_i)|$ of the number of vertices in these subgraphs. Let $\mathrm{bc}(G)$ be the smallest weight of a biclique covering of G. Let K_n be a complete graph on n vertices.

THEOREM 3.3. *If n is a power of two, then* $\mathrm{bc}(K_n) = n \log_2 n$.

PROOF. Let $n = 2^m$. We can construct a covering of K_n as follows. Assign to each vertex v its *own* vector $x_v \in \{0,1\}^m$, and consider $m = \log_2 n$ bipartite cliques $H_1, \dots, H_m$, where two vertices u and v are adjacent in H_i iff $x_u(i) = 0$ and $x_v(i) = 1$. Since every two distinct vectors must differ in at least one coordinate, each edge of K_n belongs to at least one of these bipartite cliques. Moreover, each of the cliques has weight $(n/2) + (n/2) = n$, since exactly $2^{m-1} = n/2$ of the vectors in $\{0,1\}^m$ have the same value in the i-th coordinate. So, the total weight of this covering is $mn = n \log_2 n$.

To prove the lower bound we use a probabilistic argument. Let $A_1 \times B_1, \dots, A_t \times B_t$ be a covering of K_n by bipartite cliques. For a vertex v, let m_v be the number of these cliques containing v. By the double-counting principle,

$$\sum_{i=1}^{t} (|A_i| + |B_i|) = \sum_{v=1}^{n} m_v$$

is the weight of the covering. So, it is enough to show that the right-hand sum is at least $n \log_2 n$.

To do this, we throw a fair 0-1 coin for each of the cliques $A_i \times B_i$ and remove all vertices in A_i from the graph if the outcome is 0; if the outcome is 1, then we remove B_i. Let $X = X_1 + \cdots + X_n$, where X_v is the indicator variable for the event "the vertex v survives."

Since any two vertices of K_n are joined by an edge, and since this edge is covered by at least one of the cliques, at most one vertex can survive at the end. This implies that $\mathrm{E}[X] \le 1$. On the other hand, each vertex v will survive with probability 2^{-m_v}: there are m_v steps that are "dangerous" for v, and in each of these steps the vertex v will survive with probability $1/2$. By the linearity of expectation,

$$\sum_{v=1}^{n} 2^{-m_v} = \sum_{v=1}^{n} \Pr[v \text{ survives}] = \sum_{v=1}^{n} \mathrm{E}[X_v] = \mathrm{E}[X] \le 1 \,.$$

We already know that the arithmetic mean of numbers $a_1, \dots, a_n$ is at least their geometric mean:

$$\frac{1}{n} \sum_{v=1}^{n} a_v \ge \Big(\prod_{v=1}^{n} a_v \Big)^{1/n} .$$

When applied with $a_v = 2^{-m_v}$, this yields

$$\frac{1}{n} \ge \frac{1}{n} \sum_{v=1}^{n} 2^{-m_v} \ge \Big(\prod_{v=1}^{n} 2^{-m_v} \Big)^{1/n} = 2^{-\frac{1}{n} \sum_{v=1}^{n} m_v} ,$$

from which $2^{\frac{1}{n} \sum_{v=1}^{n} m_v} \ge n$, and hence, also $\sum_{v=1}^{n} m_v \ge n \log_2 n$ follows. □

5. 2-colorable families

Let $\mathcal{F}$ be a family of subsets of some finite set. Can we color the elements of the underlying set in red and blue so that no member of $\mathcal{F}$ will be monochromatic? Such families are called 2-*colorable*.

Recall that a family is *k-uniform* if each member has exactly k elements.

THEOREM 3.4 (Erdős 1963b). *Every k-uniform family with fewer than 2^{k-1} members is 2-colorable.*

PROOF. Let $\mathcal{F}$ be an arbitrary k-uniform family of subsets of some finite set X. Consider a random 2-coloring obtained by coloring each point independently either red or blue, where each color is equally likely. Informally, we have an experiment in which a fair coin is flipped to determine the color of each point. For a member $A \in \mathcal{F}$, let X_A be the indicator random variable for the event that A is monochromatic. So, $X = \sum_{A\in\mathcal{F}} X_A$ is the total number of monochromatic members.

For a member A to be monochromatic, all its $|A| = k$ points must receive the same color. Since the colors are assigned at random and independently, this implies that each member of $\mathcal{F}$ will be monochromatic with probability at most $2 \cdot 2^{-k} = 2^{1-k}$ (factor 2 comes since we have two colors). Hence,

$$\mathrm{E}[X] = \sum_{A\in\mathcal{F}} \mathrm{E}[X_A] = \sum_{A\in\mathcal{F}} 2^{1-k} = |\mathcal{F}| \cdot 2^{1-k}.$$

Since points in our probability space are 2-colorings, the pigeonhole property of expectation implies that a coloring, leaving at most $|\mathcal{F}| \cdot 2^{1-k}$ members of $\mathcal{F}$ monochromatic, must exist.

In particular, if $|\mathcal{F}| < 2^{k-1}$ then no member of $\mathcal{F}$ will be left monochromatic. □

The proof was quite easy. So one could ask whether we can replace 2^{k-1} by, say, 4^k? By turning the probabilistic argument "on its head" it can be shown that this is *not* possible. The sets now become random and each coloring defines an event.

THEOREM 3.5 (Erdős 1964a). *If k is sufficiently large, then there exists a k-uniform family $\mathcal{F}$ such that $|\mathcal{F}| \le k^2 2^k$ and $\mathcal{F}$ is not 2-colorable.*

PROOF. Set $r = \lfloor k^2/2 \rfloor$. Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots$ be independent random members of $\binom{[r]}{k}$, that is, $\boldsymbol{A}_i$ ranges over the set of all $A \subseteq \{1, \ldots, r\}$ with $|A| = k$, and $\Pr[\boldsymbol{A}_i = A] = \binom{r}{k}^{-1}$. Consider the family $\boldsymbol{\mathcal{F}} = \{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_b\}$, where b is a parameter to be specified later. Let χ be a coloring of $\{1, \ldots, r\}$ in red and blue, with a red points and $r - a$ blue points. Using Jensen's inequality, for any such coloring and any i, we have

$$\begin{aligned}\Pr[\boldsymbol{A}_i \text{ is monochromatic}] &= \Pr[\boldsymbol{A}_i \text{ is red}] + \Pr[\boldsymbol{A}_i \text{ is blue}] \\ &= \frac{\binom{a}{k} + \binom{r-a}{k}}{\binom{r}{k}} \ge 2\binom{r/2}{k} \Big/ \binom{r}{k} := p,\end{aligned}$$

(See the next two pages for Jensen's inequality & the aymptotic formula for binomial coefficients.)

where, by the asymptotic formula for the binomial coefficients, p is about $\mathrm{e}^{-1}2^{1-k}$. Since the members $\boldsymbol{A}_i$ of $\boldsymbol{\mathcal{F}}$ are independent, the probability that a given coloring χ is legal for $\boldsymbol{\mathcal{F}}$ equals

$$\prod_{i=1}^{b} (1 - \Pr[\boldsymbol{A}_i \text{ is monochromatic}]) \le (1-p)^b.$$

Hence, the probability that at least one of all 2^r possible colorings will be legal for $\boldsymbol{\mathcal{F}}$ does not exceed $2^r(1-p)^b < \mathrm{e}^{r\ln 2 - pb}$, which is less than 1 for $b = (r \ln 2)/p = (1 + o(1))k^2 2^{k-2}\mathrm{e}\ln 2$. But this means that there must be at least one realization of the (random) family $\boldsymbol{\mathcal{F}}$, which has only b sets and which cannot be colored legally. □

Let $B(k)$ be the minimum possible number of sets in a k-uniform family which is not 2-colorable. We have already shown that

$$2^{k-1} \le B(k) \le k^2 2^k.$$

As for exact values of $B(k)$, only the first two $B(2) = 3$ and $B(3) = 7$ are known. The value $B(2) = 3$ is realized by the graph K_3.

There is yet another class of 2-colorable families, without any uniformity restriction.

THEOREM 3.6. *Let $\mathcal{F}$ be an arbitrary family of subsets of a finite set, each of which has at least two elements. If every two non-disjoint members of $\mathcal{F}$ share at least two common elements, then $\mathcal{F}$ is 2-colorable.*

PROOF. Let $X = \{x_1, \ldots, x_n\}$ be the underlying set. We will color the points $x_1, \ldots, x_n$ one-by-one so that we do not color all points of any set in $\mathcal{F}$ with the same color. Color the first point x_1 arbitrarily. Suppose that $x_1, \ldots, x_i$ are already colored. If we cannot color the next element x_{i+1} in red then this means that there is a set $A \in \mathcal{F}$ such that $A \subseteq \{x_1, \ldots, x_{i+1}\}$, $x_{i+1} \in A$ and all the points in $A \setminus \{x_{i+1}\}$ are red. Similarly, if we cannot color the next element x_{i+1} in blue, then there is a set $B \in \mathcal{F}$ such that $B \subseteq \{x_1, \ldots, x_{i+1}\}$, $x_{i+1} \in B$ and all the points in $B \setminus \{x_{i+1}\}$ are blue. But then $A \cap B = \{x_{i+1}\}$, a contradiction. Thus, we *can* color the point x_{i+1} either red or blue. Proceeding in this way we will finally color all the points and no set of $\mathcal{F}$ becomes monochromatic. □

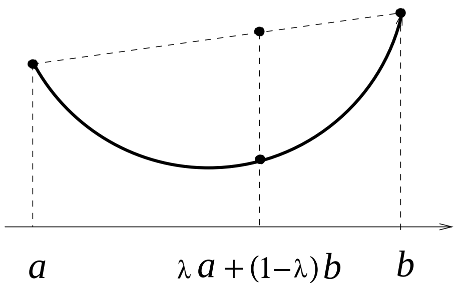

FIGURE 1. A convex function.

We mention one important inequality, which is especially useful when dealing with averages. A real-valued function $f(x)$ is *convex* if

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b)\,,$$

for any $0 \leq \lambda \leq 1$. From a geometrical point of view, the convexity of f means that if we draw a line l through points $(a, f(a))$ and $(b, f(b))$, then the graph of the curve $f(z)$ must lie below that of $l(z)$ for $z \in [a, b]$. Thus, for a function f to be convex it is sufficient that its second derivative is nonnegative.

PROPOSITION 1.12 (Jensen's Inequality). *If $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^{n} \lambda_i = 1$ and f is convex, then*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)\,. \tag{14}$$

PROOF. Easy induction on the number of summands n. For $n = 2$ this is true, so assume the inequality holds for the number of summands up to n, and prove it for $n+1$. For this it is enough to replace the sum of the first two terms in $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{n+1} x_{n+1}$ by the term

$$(\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2\right)\,,$$

and apply the induction hypothesis. □

Tighter (asymptotic) estimates for binomial coefficient can be obtained using the famous *Stirling formula* for the factorial:

$$n! = \left(\frac{n}{\mathrm{e}}\right)^n \sqrt{2\pi n}\,\mathrm{e}^{\alpha_n}\,, \tag{7}$$

where $1/(12n+1) < \alpha_n < 1/12n$. This leads, for example, to the following elementary but very useful asymptotic formula for the k-th factorial:

$$(n)_k = n^k \mathrm{e}^{-\frac{k^2}{2n} - \frac{k^3}{6n^2} + o(1)} \quad \text{valid for } k = o(n^{3/4}), \tag{8}$$

and hence, for binomial coefficients:

$$\binom{n}{k} = \frac{n^k \mathrm{e}^{-\frac{k^2}{2n} - \frac{k^3}{6n^2}}}{k!}(1 + o(1))\,. \tag{9}$$